ROSALES, OSBIN FRENET PERDOMO, M.S. Measuring Distances Between Nucleotides Using CW- EPR and Dual Site-Directed Spin Labels to Distinguish Parallel Quadruplexes from B-DNA. (2018)
Directed by Dr. Ethan W. Taylor 42 pp.

The proper identification of DNA secondary structure is paramount for determining possible factors that can influence rates of genomic transcription. Structures such as DNA quadruplexes (QPX), which are found in guanine-rich regions of DNA, can have a significant effect on an organism, due to their ability to influence transcription rates under certain conditions. Aside from NMR and X–ray crystallography, Circular Dichroism (CD) has been shown to be an easy and effective method to distinguish between classical double helical A, B, and Z DNA, and other secondary structures. However, complications have arisen from this technique, due to the similarities in B DNA and parallel QPX spectra, because both have positive CD peaks at 260 nm, negative at 245 nm and a positive peak around 210 nm and 205 nm, respectively. Currently, a G-C rich duplex DNA oligonucleotide, G21, which contains sequences found in the promoter region of the ChAT gene, has shown both B-DNA and parallel quadruplex spectra. However, under different salts and salt concentrations, changes of peak intensities were observed. This could be due to the formation of different secondary structures that are not distinguishable under CD. To address this, a continuous wave electron paramagnetic resonance (CW-EPR) technique was used. This technique utilizes a pair of site-directed spin labels (SDSL) attached via phosphorothioate bonds to two different modified nucleotides, to determine the distances between them. Variations in

the measured distance will be used to distinguish between B-DNA or parallel QPX, as either structure could be possible from the CD spectrum alone.

MEASURING DISTANCES BETWEEN NUCLEOTIDES USING CW-EPR AND DUAL SITE- DIRECTED SPIN LABELS TO DISTINGUISH PARALLEL QUADRUPLEXES FROM B-DNA

by

Osbin Frenet Perdomo Rosales

A Thesis Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Greensboro
2018

Approved by

_____
Committee Chair

APPROVAL PAGE

This thesis written by OSBIN FRENET PERDOMO ROSALES has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____

Committee Members _____

_____

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

1.1 Review of Literature

1.1.1 Choline Acetyltransferase

Acetylcholine (Ach) is an important neurotransmitter that aids in nerve cell to cell communication.[1] Its bioregulation is very dependent on the enzyme responsible for its biosynthesis, choline acetyltransferase (ChAT).[1]ChAT uses the substrates choline and acetyl-CoA, to make acetyl choline.[1]The source of choline can come from either choline downstream biosynthetic pathways or from ACh that has been converted to choline by cholinesterases in the synaptic cleft.[1]

Lower concentrations and activity of ChAT can cause complication and has been associated with schizophrenia, Alzheimer's disease, and other neurological disorders.[2] An increase in acetylcholine can also have negative effects as seen in the many acetylcholine receptor agonist such as nicotine that have various adverse effects on the body.[3] The genetic expression of ChAT is therefore very important. Like other enzymes ChAT is expressed from the ChAT gene.[1] The ChAT gene, like all others, is composed of three sections, promotor region, the coding region, and the transcription termination site. The promoter site which signals the genetic machinery, such as helicases that unwind DNA from two strands to one strand necessary to initiate transcription, is very important for gene expression.[4]

1.1.2 DNA Quadruplexes

The secondary structure of DNA can heavily influence gene expression.[4] The most common secondary structure that DNA can have is the classical double helical structure that was elucidated by Watson and Crick known as B-DNA. These structures are caused by the base pairing of thymine (T) to adenosine (A), and cytosine (C) to guanine (G). Other structures based on the classical base pairing are A and Z DNA which vary from B- DNA based on differences between ribose ring puckering and bond orientation for example.[5] However, in rich guanine environments Hoogsten hydrogen bonding can form between the guanine bases forming planar tetrads, figure 1.[6]

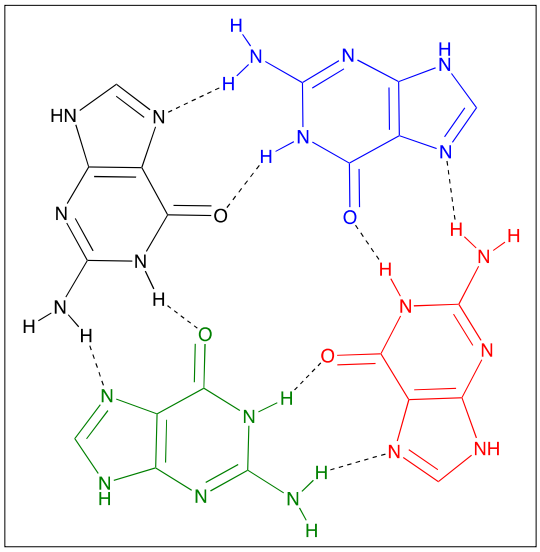

Figure 1. DNA Tetraplex Formed by Four Guanine Bases. Four guanine residues come together via Hoogsten hydrogen bonds, depicted here in the dashed lines, to create a planar tetrad. The tetrads are major components of G-QPX.

When two or more tetrads are present, they stack on top of each other. The structure that is created by the stacking of the guanine tetrads are known as DNA Quadruplexes (G-QPX).[6] The tetrads are held together by a monovalent cation, figure 2.[7] The most common monovalent cations were found to stabilize G-QPX are K^+, Na^+, and Li^+, in order of most to least stable.[7] There are however G-QPX in the literature that have specificity to just one monovalent cation, as it is in Yang *et al*, who report a K^+ specific quadruplex used to measure potassium in human blood.[8]

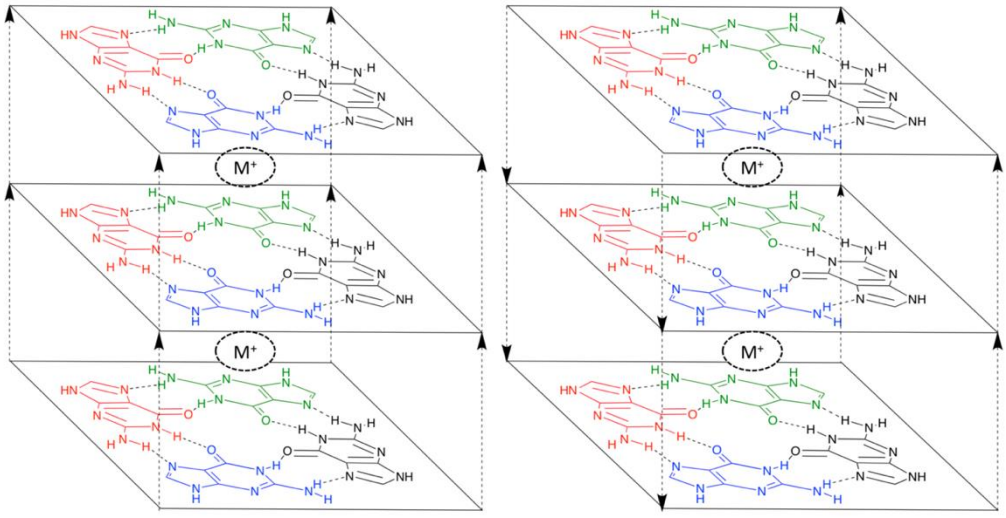

Figure 2. Parallel and Anti-parallel G-QPX. DNA G-QPX formed by the stacking of guanine tetrads held together by a monovalent cation. The colors represent different G runs demonstrating the contribution each G run has to the overall G-QPX structure. The direction of the G runs relative to each other determines whether a G-QPX is parallel (left) or anti-parallel (right).

G-QPX are formed by four guanine repeats that contain 2 or more consecutive guanines.[9] Each guanine repeat contribute one guanine to each tetrad and are separated by 1-9 nucleotides.[9] The orientation of the G runs to each other in the G-QPX determines whether the G-QPX is parallel, antiparallel, or both.[7] The number of tetrads the four guanine repeats can make depends on the lowest number of consecutive guanine found in any of the four repeats.[7] In biological systems, the G-QPX guanine repeats have been found in a single strand.[6] However, synthetic oligonucleotides made with telomeric repeats of d(TTAGGG) have been show *in vitro* using NMR and X-ray crystallography to form quadruplexes.[6]

1.1.3 G-QPX in the Human Genome and Drug Targets

The number of possible G-QPX forming sequences in the human genome is staggering. The classical estimate shows that there are around ~ 370,000 potential G-QPX forming sequences.[10] Recent studies however show that there can be over 700,000 G-QPX forming sequences.[9,11]. The studies also showed that most of the G-QPX forming sequences were near promoter regions and many oncogenes.[4,11] DNA G-QPX have been reported to influence transcription.[4] G-QPX can increase transcription by forming on one strand increasing transcription on the complementary strand.[4] The formation of G-QPX can decrease transcription by obstructing transcription on the G-QPX forming strand.[4] G-QPX can also aid in the recruitment of proteins that can either increase or decrease transcription.[4]

This prevalence and the ability for G-QPX to regulate gene expression makes them a subject of research interest as potential drug targets. Although there is no drug in

the market that targets G-QPX, there have been some candidates.[12] One of the candidates as a G-QPX targeting drug is berberine, figure 3.[13] Berberine has been found to inhibit telomerase by stabilizing telomeric G-QPX.[13] Telomerase is an enzyme that extends the telomeric ends of DNA and has been shown to be upregulated in cancer and is associated with cellular immortality.[12,13]

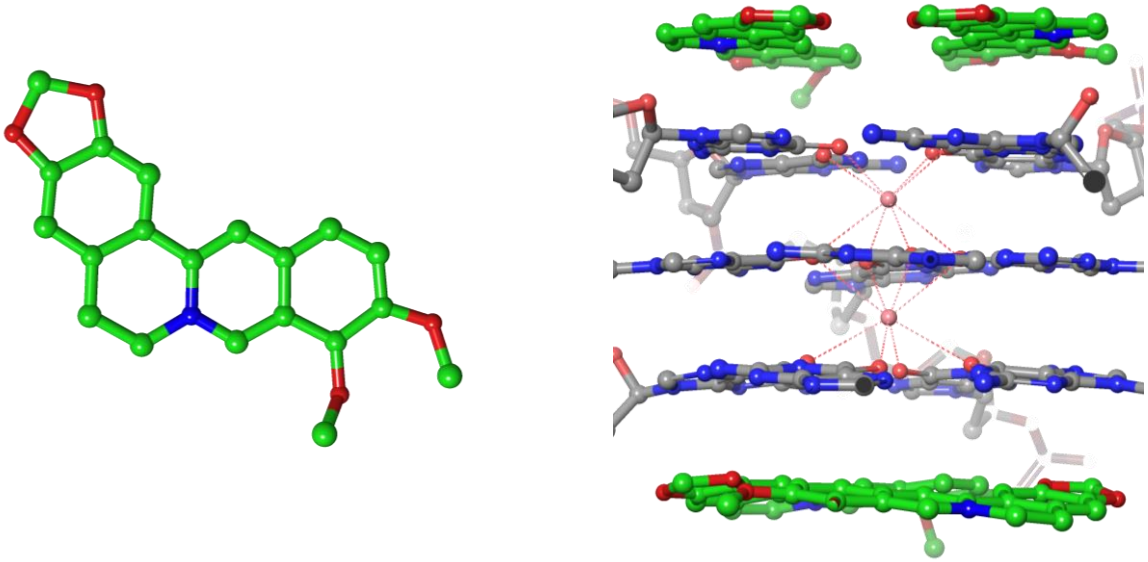

Figure 3. Berberine Structure and Berberine Molecules Complexed to Telomeric G-QPX.[13] Berberine (left) demonstrates the planar structure of most G-QPX targeting molecules.[13] A telomeric G-QPX is being stabilized by berberine that stacks above and below the G-QPX.[13]

1.2 Previous Data

Due to G-QPX prevalence in the human genome, a bioinformatics study was conducted in the Taylor group. The project looked at genes in the nervous system which

had not been studied for the presence of GQPX using G-QPX forming prediction and stability programs.[14] Overall three programs were used, QGRS mapper, G4 calculator, and Quadparser.[14] The programs looked at possible G-QPX forming DNA sequences using parameters such as number of Gs in the G runs, and the number of nucleotides between the G runs.[14] These programs also determined the stability of the G-QPX by using parameters such as evenness of number of nucleotides between G runs and the density of G in each run.[14]

5'-GGG −GAT-GCC-GCC-CGG-GGG-AGC-CTG-AGG-GAC-CCG-CTC-CAG-CTA-GGC-ACG-CCC-CCG-CCC-TTT-GAG-GAC-ACG-CCC-CAC-ACC-AGC-CTC-AGA-GCT-CTG-A-3'

3'-CCC-CTA-CGG-CGG-GCC-CCC-TCG-GAC-TCC-TGG-GCG-AGG-TCG-ATC-CGT-GCG-GTC-CGG-GGC-GGG-AAA-CTC-CTG-TGC-GGG-GTG-TGG-TCG-GAG-TCT-CGA-GAC-T-5'

Figure 4. 96 BP Sequence Containing Possible G-QPX Within the Active Promoter Region of the ChAT Gene. The positive strand (top) contains one G-QPX sequence highlighted in blue.[14] The negative strand (bottom) contain three possible G-QPX sequences.[14] This active promoter sequence was found using the transcription element search system and the PROMO database.[14]

This bioinformatics project also looked at G-QPX in the nervous system by looking in the active promoter region of the ChAT gene.[14] The active promoter region was determined by looking at reported transcription factor binding sites (TFBS) using the transcription element search system and the PROMO database.[14] An active promoter region 97 base pairs (bp), figure 4, was found to contain G-QPX forming sequences in

both the positive and the negative strand.[14] The region was found to contain one possible G-QPX sequence, G29, in the positive strand, and 3 possible G-QPX sequences, G17, G17-2, and G30, in the negative strand, table 1.[14] Due to the unusual overlapping G-QPX of G30 on the negative strand and G29 on the positive strand, it was hypothesized that a bimolecular quadruplex could form.[14] This resulted in the analyzing of ChAT G21 which consisted of G runs of both ChAT G30 and G29.[14]

Table 1. Possible G-QPX Sequences Found in the 96 BP of the Active Promoter Region of the ChAT Gene. The G-GPX sequences were found using the G-QPX predicting programs QGRS mapper, G4 calculator, and Quadparser.[14]

Name	Sequence	Strand
ChAT G17	3’ GCGGGGTGTGGTCGGAG 5’	Negative
ChAT G17-2	3’-GCGGTCCGGGGCGGGAA-5’	Negative
ChAT G30	3’-GCGGGCCCCCTCGGACTCCTGGGCGAGGTC-5’	Negative
ChAT G29	5’-GGGGATGCCGCCCGGGGGAGCCTGAGGAC-3’	Positive
ChAT G21	5’-CCCGGGGGGAGCCTGAGGACCC-3’	Positive
ChAT G21	3’-GGGCCCCCTCGGACTCCTGGG-5’	Negative

1.2.1 Biological Data

In order to verify the results of the bioinformatics project, biological assays were conducted. These assays used a modified wild type GFP reporter gene plasmid that was modified to include the active ChAT promoter region.[14] This plasmid was transfected into human neuroblastoma cell line.[14] Along with growth factors, aconitine, and 5,10,15,20-tetrakis(N-methylpyridinium-4-yl)-21,23H-porphyrin (TMPyP4) were also

used.[14] TMPyP4 was used because it is a well-known G-QPX-stabilizer drug, due to its planar structure and its electropositivity.[15]

Since ChAT is essential in nerve cells which undergo an action potential that is caused by the intake of Na^+ , aconitine, a sodium gate channel opening drug, was used to simulate an action potential.[14] The reason for the simulation of the action potential was to determine if a specificity for K^+ was present in the G-QPX found in the active promoter of the region of the ChAT gene.[14] The specificity of K^+ could potentially melt the G-QPX due to the displaced K^+ by the high amounts of Na^+ .[14] The melting of the G-QPX in the active promoter region could cause an increase in expression of the ChAT gene producing ACh that was depleted during the action potential.[14] The biological assays determined that when the cells introduced to TMPyP4, a decrease in the production of GFP was seen.[14] When aconitine was used however an increase in production of GFP was seen but lower than wild type GFP.[14] The biological assay showed that one or multiple G-QPX were present in this active promoter region and also suggested that the G-QPX were functionally inhibited under high concentrations of Na^+.[14]

1.2.2 Physical Data

Circular dichroism (CD) analysis was also done on each individual sequence.[14] CD spectroscopy works by using the difference between left and right circular polarized light at different wavelengths to determine structural changes in chiral molecules.[16] Since DNA is a chiral biopolymer, CD can be used as well as it also allows DNA to be studied under various conditions and solvents.[16] Although this method does not elucidate the structure of a macromolecule, it can help distinguish between secondary structures.[16] The

CD analysis was done using 10 to 100 mM concentrations of Na^+, Li^+, and K^+.[14] These concentration are more than sufficient as G-QPX have been shown to form in concentration of 0.5 mM to 10 mM K^+ in a biological system.[8]

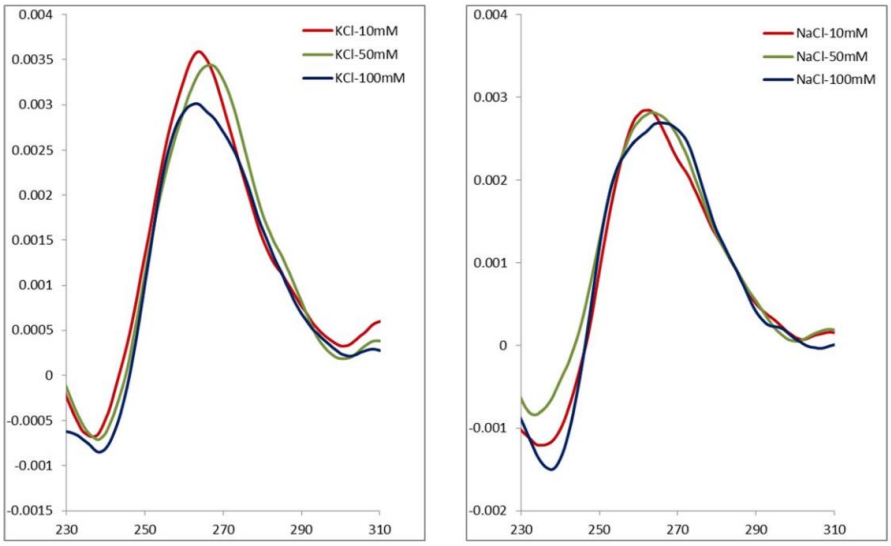

Figure 5. CD Spectra of ChAT G21 Sequence Using Various Concentrations of NaCl and KCl. The CD spectra of ChAT G21 under high amounts of KCl (left) showed a positive peak at 260 nm and negative peak at 245 nm which indicated the presence of a parallel G-QPX. The CD spectra of ChAT G21 under high amounts of NaCl (Right) showed several peaks between 260 nm and 280 nm and negative peak at 245 nm which indicated the presence of a B-DNA. The X axis represents wavelength in nm and the y axis represents the ellipticity.

The concentration range of 10 to 100 mM for the monovalent cation is also biologically relevant as the concentration of potassium outside of the cell is approximately 5 mM and the concentration of Na^+ is approximately 145 mM.[8]The CD

analysis revealed that a G-QPX could be present in ChAT G29 and ChAT G30.[14] This was unusual due to the proximity of the sequences, as ChAT G29 belonged to the positive strand and ChAT G30 was on the negative strand, and thus both sequences contained complementary bases.[14] Due to the proximity of the G-QPX sequences, ChAT G21 was analyzed as a duplex containing sequences from both G29 and G30.[14] ChAT G21 CD spectra showed that a parallel G-QPX could be present, figure 5.[14] Since the spectra of parallel G-QPX and B-DNA is very similar, future studies were needed to determine which structure was present.

1.3 Central Hypothesis and Objectives

If a G-QPX is present in this region, this would be the first time that a functional G-QPX will be found in a nervous system gene.[14] This would also be the first time that a "biomolecular" G-QPX consisting of a DNA sequence and its complementary strand has ever been discovered.[6,10] Although there was already CD data for G21, this project hoped to expand this data. In particular, the ambiguity between the B-DNA CD spectra and parallel G-QPX needed to be resolved. It was hoped that the ambiguity would be resolved by using B-DNA and parallel G-QPX controls, and by looking at wavelength ranges not previously studied for ChAT G21 to find structure distinguishing features. This project will also take a closer look into the K^+ specificity observed in the biological assays by testing ChAT G21 in varying concentration of Na^+ and K^+.

The main purpose of this project is to investigate ChAT G 21 duplex using other physical methods aside from CD to determine its structure. This is going to be done by using continuous wave electron paramagnetic resonance (EPR). With EPR, nucleotides

could be specifically tagged with a nitroxide radical.[17] The distances between two nitroxide radicals then can be measured for up to distance of 25 A.[18] Also, how fast ChAT G21 moves in solution, the rotational correlation time, can also be measured. Both the changes in the rotational correlation time and distances can then be used to determine if a G-QPX or B-DNA is present in ChAT G21.

CHAPTER II

METHODOLOGY

2.1. Methods to Analyze DNA Structures

There are several methods to analyze G-QPX and obtain structural data. Among them are nuclear magnetic resonance (NMR),[7] X-ray crystallography,[13] EPR[19], and CD[5,20]. There are advantages and disadvantages to each of them. To fully elucidate the structure of a macromolecule, NMR and X-ray crystallography are used. There are however disadvantages to both methods. NMR requires a high amount of sample due to the large size of the oligonucleotides and can be very costly.[7] This is further complicated with the high level of purity needed to analyze the sample. In order to distinguish between nucleotides and determine the distance between intramolecular and intermolecular bonds, ^{14}N and ^{13}C, tagged nucleotides are needed adding to the complexity.[7] X-ray crystallography has its setbacks, as not all compounds can be crystallized and also due to the need for access to X-ray crystallography facilities.

2.2. Circular Dichroism

Circular dichroism is an inexpensive method and requires as little as 20 μg / mL of material.[20] Although CD cannot elucidate the structure of DNA, it can be used to distinguish between secondary structures.[5,16] The differences between the secondary structures come from the differences between positive and negative peaks in the 200 to

300 nm range. The differences between the secondary structures can be attributed to many factors. The puckering of the sugar ring between B-DNA and A-DNA causes major structural changes.[21] This leads to a drastic change in spectra since B-DNA shows a positive peak or peaks between 260 to 280 nm, a negative peak at 245 nm and a positive peak at 205.[16] A-DNA shows a positive peak at 260 nm and a negative a 210 nm. For G-QPX the major difference in the CD spectra is due to the orientation of the DNA strands and the orientation of the glycosidic bond between the sugar and base.[5] Parallel G-QPX have a positive peak at 260 and 210 nm as well as a negative peak at 245 nm.[16] Antiparallel quadruplexes share a positive peak at 290 and 210 nm and negative peaks at 260 and 245 nm.[16] These features in the CD spectra allow them to be distinguished from each other.

Table 2. Circular Dichroism Peaks of Relevant DNA Secondary Structures. The positive peaks in the CD spectra are demonstrated as (+) and the negative peaks are (-). The data in this table was comprised from Vorlíčková et al, and Kypr et al.[5,16]

G-QPX Antiparallel	G-QPX Parallel	B-DNA	Z-DNA	A-DNA
290 nm (+)	260 nm (+)	280-260 nm (+) (Peak, Peaks or Stretch)	290 nm (+)	260 nm (+)
260 nm (-)	245 nm (-)	245 nm (-)	260 nm (-)	210 nm (-)
245 nm (-)	210 nm (+)	205 nm (+)	205 nm (-)	
210 nm (+)				

When it comes to analyzing duplex DNA, issues can arise. The positive peak or peaks of B-DNA between 260-280 nm of B-DNA can vary greatly sometimes resembling

a single 260 nm peak similar to those shown by parallel G-QPX.[16] This leads to a great disadvantage since the only distinguishing feature between B-DNA and a parallel G-QPX is the negative peaks at 210 and 205 nm, which are in practice often indistinguishable.[16] Therefore, for this project this range was investigated more closely as well as the use of B-DNA and parallel G-QPX controls. This was expected to yield more evidence to the actual structure of ChAT G21.

2.3 Electron Paramagnetic Resonance

EPR is another method that can provide important structural data. This method, just like CD, cannot fully elucidate the structure of macromolecules, but distance measurements between nucleotides[18] and changes in mobility can be obtained[22,23]. Unlike NMR that uses radiowaves in a varying magnetic field, EPR uses microwaves to characterize unpaired electrons and the system that they are confined in. This is one of the major advantages to EPR since only specific molecules such as organic radicals and metals can be detected. Therefore, all other molecules do not cause unwanted interference. Another key advantage of EPR is its higher sensitivity compared to NMR. This and the unwanted interference from unwanted molecules results in smaller sample concentration and the use of other none deuterated solvents and molecules.

2.3.1 Site Directed Spin Labeling

Since DNA does not have any unpaired electrons or radicals, a spin label would have to be attached. The most common types of spin labels that are used to analyze macromolecules are nitroxide radicals.[24] Nitroxide radicals are used because of radical stability and their ability to be modified using organic synthesis.[17,24–27] The nomenclature

of nitroxide radical spin labels is very broad, but the moiety seen in all of them is the di-tert-butyl nitroxide moiety that is housed either in a six or five-member ring.[24] For this project the nitroxide radical 3-(iodomethyl)-2,2,5,5-tetramethyl-3-pyrrolin-1-yloxy.

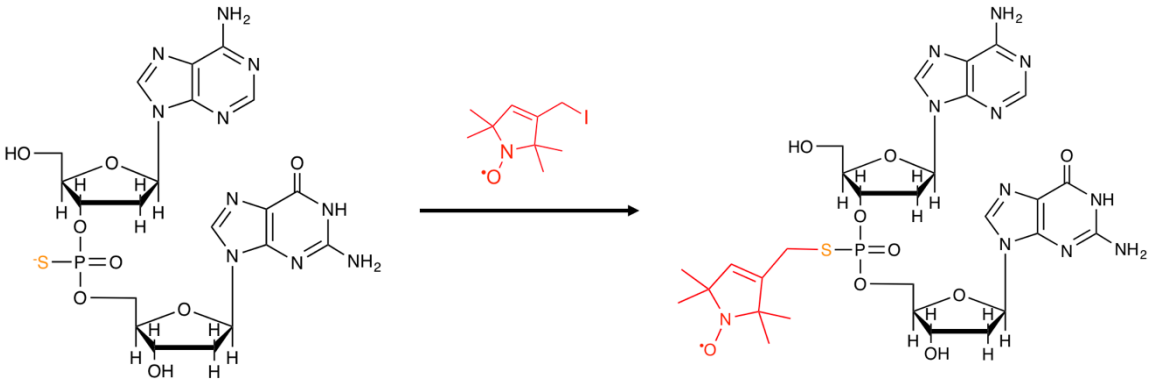

Figure 6. Phosphorothioate Modified Oligonucleotide Spin Labeling Using a Nitroxide Radical. In a simple substitution reaction, the nitroxide radical (red) can be added to the phosphorothioate modification (orange) in the oligonucleotides.[17] The phosphorothioate modification is amongst the least intrusive since the intermolecular interactions occur in the bases.

DNA has no radicals and unpaired electrons but a radical or spin label can be attached at desired sites via structurally modified nucleotides. Although there are many different modifications that can be done on nucleotides, the least intrusive and most common modification to nucleotides in which a spin label can be easily attached is the phosphorothioate modification.[17] The phosphorothioate modification is the modification that involves the replacement of a nonbonding oxygen in the phosphate DNA backbone to a sulfur atom. In the presence of a strong electrophile, the sulfur atom could act as a

nucleophile, joining the spin label to the DNA molecule. In the presence of a phosphorothioate modified nucleotide the nitroxide radical used for this project, 3-(iodomethyl)-2,2,5,5-tetramethyl-3-pyrrolin-1-yloxy, has a labeling efficiency of nearly 100%, figure 6.[17] The only subsequent step before analysis in the EPR spectrophotometer is the purification of the tagged oligonucleotide to ensure no untagged spin labels are present.[17] The purification could either be done via high performance liquid chromatography or SDS-PAGE.[17]

2.3.2 Mobility Measurements

How fast a molecule is moving in solution can be determined using EPR spectroscopy. When a molecule is moving very slowly the EPR spectra is anisotropic. However when the molecule is moving fast the EPR signal are averaged out across all the x,y, and z axis. [22] Most mobility analysis of macromolecules are done in the fast isotropic region. For nitroxide radicals three peaks are produced by the lone electron in the oxygen atom and by the nuclear splitting of ^{14}N, in the isotropic region. Most mobility analysis for macromolecules has been done in this region and it has been observed that faster the radical moves, the narrower the three EPR spectra peaks becomes.[22] When the molecule slows down in this region the three EPR spectra peaks becomes broader.[22] In an oligonucleotide with a nitroxide radical firmly attached to it, the interspin vector rotational correlation time of the radical equals the rotational correlation time (τ) of the DNA molecule.[23]

Once the EPR spectrum of the DNA molecule is obtained, the rotational correlation time obtained using the EPR spectra simulation program Easyspin.[22] Easyspin

can simulate the nitroxide radical spin system with specific parameters such as the rotational correlation time.[22] Once the spectrum is generated, the theoretical spectrum can be compared with the experimental spectrum to approximate the correlation rotational time. The changes in rotational correlation time have been seen between G-QPX and single stranded form of DNA. In Okamoto et al, it was reported that a G-QPX with a nitroxide radical had a rotational correlation time of 0.32 ns and the non G-QPX single stranded form had a rotational correlation time of 1.56 ns.[23] The changes in rotational correlation time was attributed to the compact structure of G-QPX allowing for more movement in solution compared to the non-compact structures seen in a non G-QPX single stranded form molecule.[23]

2.3.3 Distance Measurements

When two spin labels are attached to different phosphorothioate modified nucleotides in an oligonucleotide, distance measurements for up to distances of up to 25 A can be obtained.[18,29–31] Although there are two methods to determine the distances between the nitroxide radical, the most convenient method is the high mobility method.[30,31] This method can be used for the study of ChAT G21 because of its molecular mass is less than 20 kDa.[30,31] The low molecular mass of the oligonucleotide allows for a high rotation correlation with the anisotropic EPR spectra of the nitroxide radical to be averaged out.[30,31]The high mobility distance measurement method requires the rotational correlation time of the molecule.[30,31] For the high mobility method, once the rotational correlation time is acquired, three EPR spectra are needed.[30,31]

The first two spectra are those of each phosphorothioate modified nucleotide with the nitroxide radical attached to it.[30,31] The third spectrum is that of the DNA molecule with both nitroxide radicals attached to the phosphorothioate modified nucleotides.[30,31] The distances between the nitroxide radical (R_{dd}) can be obtained using equation 1.[30,31] Equation 1 uses γ as the gyromagnetic ratio of the electron, ν is the microwave frequency, and g is the electronic g-factor.[30,31] The spectral line broadening ΔH_{dd} is obtained from the average of the two single labeled DNA spectrum and the double labeled DNA spectra.[30,31] This spectral line broadening is measured by using the line width at the half height of each peak.[30,31]

$$\Delta H_{dd} = \frac{3}{10}\frac{\gamma^4}{g}\frac{\tau}{R_{dd}^6}\left(3 + \frac{5}{1+\nu^2\tau^2} + \frac{2}{1+4\nu^2\tau^2}\right) \qquad (1)$$

CHAPTER III

EXPERIMENTAL

3.1. Circular Dichroism

The oligonucleotides used for this project were purchased from IDT DNA and purified using reverse phase HPLC. For the CD assays, 5 μM ChAT G21 samples were measured using different concentrations of NaCl and KCl. The high KCl samples using a 20 mM potassium phosphate buffer and with varying concentrations of KCl. The high NaCl samples were made using TEA buffer consisting of 30 mM Tris HCl and 10 mM EDTA. The samples were heated up to 90 ℃ for 5 minutes and allowed to cool gradually overnight. The prepared CD samples were analyzed using a 1 mm path, 300 microliter CD cuvette in an Olis DSM 17 CD spectrophotometer. All samples were analyzed under 25 ℃. The samples were scanned from 200 nm to 300 nm at a rate of 100 nm/min with a response time of 1 s and bandwidth of 2 nm. All sample spectra were baseline corrected using their corresponding buffers.

3.2 Synthesis of Nitroxide Radical

All solvents and reagents were acquired from commercial sources and were used without additional purification. Anhydrous solvents required for some reactions were dried using the standard methods. All reactions were monitored using thin layer chromatography (TLC) using silica G F254 precoated glass-backed plates. Flash column chromatography was done using flash grade silica gel (partical size: 40-63 μm, 230 x 400

mesh). High resolution mass spectra (HRMS) were acquired on an Orbitrap XL MS system.

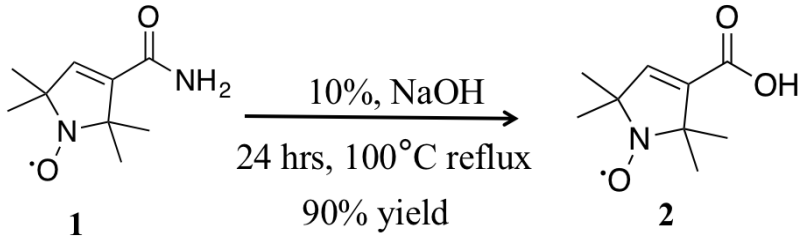

Figure 7. Compound 2.

To an aqueous solution of 10% volume per weight of NaOH, 200 mg of compound 6 was added. The reaction was allowed to reflux for 24 hours at 100 ℃ in an oil bath. The reaction was quenched to pH 2 using HCl and was extracted using dichloromethane (DCM). The organic layer was kept and dried using magnesium sulfate. This reaction resulted in 90% yield. Compound 2 retained the yellow color of compound 1 and was not purified to optimize yield. HRMS ($C_9H_{14}NO_3$•, ESI): Calculated 185.1052 $[M+H]^{+1}$, 183.0895 $[M-H]^{-1}$, found: 183.09007.

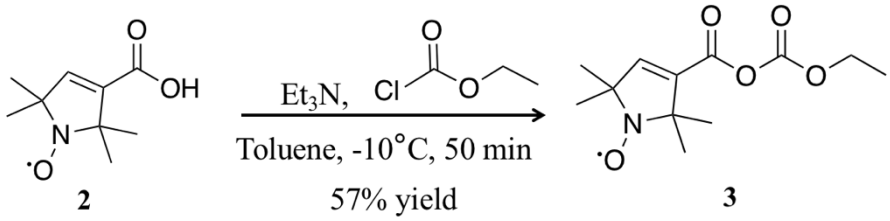

Figure 8. Compound 3.

Compound 2 was the dissolved using anhydrous toluene in a flame dried round bottom flask under argon. The reaction was cooled to -10 ℃ using a 1 to 1 ratio of ice and acetone. Once cooled 1.5 equivalents of ethyl chloroformate and 2 equivalents of dry triethylamine were added dropwise. The reaction was then allowed to sit 50 min at -10 ℃. The product was concentrated using rotary evaporation. The concentrated product was then then dissolved in diethyl ether (Et_2O). A white precipitate formed which was filtered off, washed with Et_2O and discarded. The product compound 3 was then recrystallized using hexane overnight. This reaction produced yellow crystals and resulted in 57% yield. HRMS ($C_{12}H_{18}NO_5$•, ESI): Calculated: 256.1185 $[M^{\cdot}]^{+1}$, 257.1263 $[M^{+1}H]^{+1}$, 258.1341 $[M^{+2}H]^{+1}$, found: 256.11780, 257.12571, 258.13346.

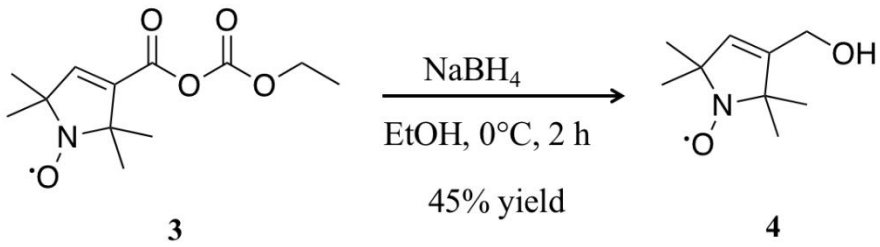

Figure 9. Compound 4.

Compound 3 was dissolved in 100% ethanol in a flame dried round bottom flask under argon. The reaction was cooled to 0 ℃ using an ice bath and 2.5 equivalents of $NaBH_4$ was added. The reduced product was recovered by concentrating the sample under high vacuum. The concentrated produced was resuspended in water. The product was then extracted using Et_2O. The product was purified using column chromatography using 9:1 Et_2O and petroleum ether. This reaction produced a white crystal as the product

and resulted in a 45% yield. An NMR analysis of the product was attempted by using Chloroform-D and adding 1:1 ratio of phenylhydrazine to quench the radical. This was unsuccessfully since compound 4 became insoluble once protonated and also due to small amounts of material. HRMS($C_9H_{16}NO_2$•, ESI): 170.1181 $[M]^{+1}$, found: 170.1175.

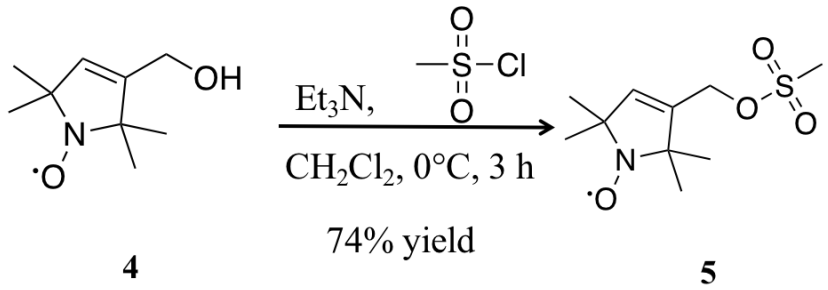

Figure 10. Compound 5.

Compound 4 was added to a flame dried round bottom flask under argon and dissolved in dry DCM. The dissolved starting material was cooled to 0 ℃ using an ice bath. Once cooled 1.2 equivalents of dry triethylamine and methylsulfonyl chloride were added dropwise. After the addition of the reagents the reaction was then brought to room temperature and allowed to react for 3 hours. The reaction was then diluted with DCM and quenched with 5% $NaHCO_3$. The organic layer was kept and dried with magnesium sulfate. This reaction was purified using column chromatography using 3:1 Et_2O and petroleum ether. This reaction resulted in 74% yield and produced a yellow product. HRMS($C_{10}H_{18}NO_4S$•, ESI): 248.0951$[M]^{+1}$, found 248.095

CHAPTER IV

RESULTS AND DISCUSSION

4.1 Circular Dichroism

As previously mentioned, both parallel quadruplexes and B-DNA have positive peaks at 210 nm and 205 nm, respectively. From the data generated, the determination of the 210 nm and 205 nm peak is not easy to determine due to their proximity, figures 11 and 12. A trend that can be easily determined is the negative peak at 245 nm. As the concentration of K^+ is increased, this peak is shown to become more negative. Also, a trend could not be determined for the 260 nm peak. This lack of trend in the 260 nm peak and the negative trend at 245 nm were both seen in the previous study.[14]

The CD data for G21 under varying Na^+ concentrations was also reexamined with greater emphasis at the below 230 nm range in the CD spectra, figure 12. The NaCl CD data provides also very little evidence towards structural differences at the different concentrations. However, the G21 spectrum in 100 mM NaCl is unique. This spectrum shows a clear deviation from all the other spectra. It shows a significant drop in the height of the 260 nm peak and also a huge increase in the 240 nm negative peak. This could entail a structural change in the DNA from a parallel quadruplex to B-DNA. The high amounts of Na^+ could have displaced monovalent cations that were present in the oligonucleotide that were left over from its synthesis.

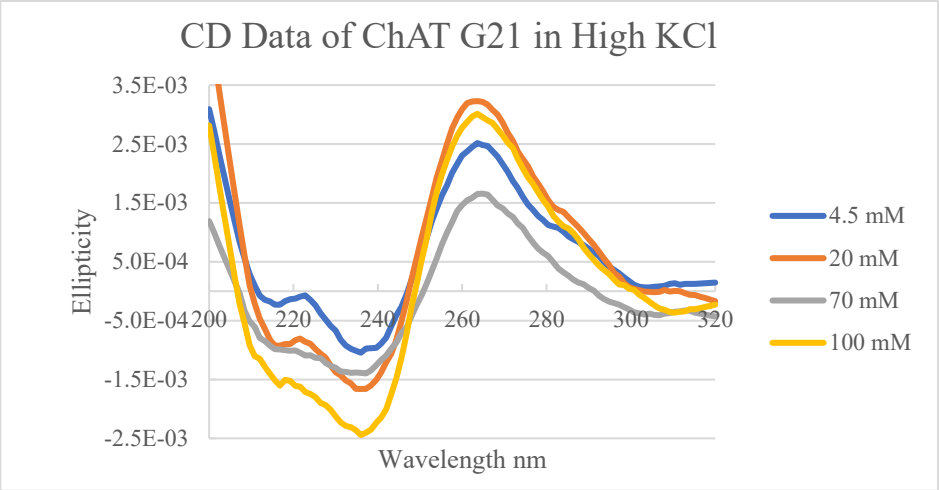

Figure 11. CD Data of ChAT G21 in High KCl. The CD data of G21 from 200 nm to 320 nm is shown here. The data in this graph rules out other DNA structures but not B-DNA and parallel quadruplexes.

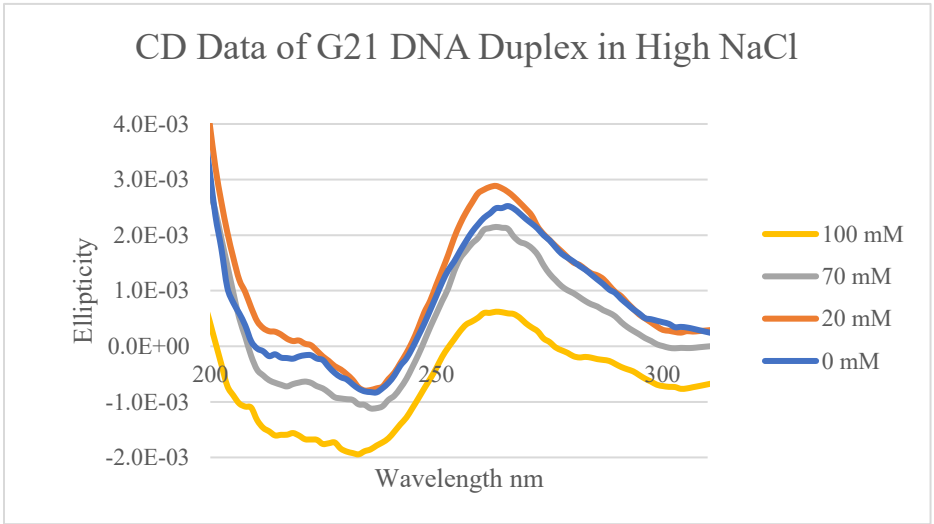

Figure 12. CD Data of ChAT G21 in High NaCl. The CD data of G21 from 200 nm to 310 nm at different concentrations of NaCl are shown here. The results from this graph are inconclusive and need to be repeated as the samples could have contained residual K+.

4.2 Future CD work

Although the CD spectra of ChAT G21 in high Na^+ and K^+ were inconclusive, there is further work that can be done. Controls could be run for both B-DNA and parallel quadruplexes. For the B-DNA control G21m can be used, table 3. G21m has the same G:C:T:A ratio as ChAT G21 but in a different sequence. The randomization of ChAT G21 will allow for the G runs to be separated and for a G-QPX to form. The spectra can then be compared to the high NaCl and KCl ChAT G21 CD spectra.

Table 3. G21m Sequence. This sequence was made from ChAT G21m and contains the same G:C:T:A ratio.

G21m	5′-GCCGACGAGTGCACGCGCGCG-3′
	3′-CGGCTGCTCACGTGCGCGCGC-5′

For the parallel G-QPX control, a DNA sequence with its structure elucidated via NMR or X-Ray crystallography and with a CD spectrum can be used. One of the DNA sequences that meets these requirements is Pu30 (5' - AGG GGC GGG CGC GGG AGG AAG GGG GCG GGA-3').[32] Pu30 is ideal due to its length of 30 nucleotides resembling ChAT G21 even further. The Pu30 CD spectrum also has a positive peak at 260 nm and a negative peak at 245 nm, there is however no data in the 200 to 230 nm range.[32] The 210 nm positive peak is missing since the CD spectra that was reported was from 230 nm to 320 nm.[32] Just like G21m, Pu30 could be compared to the ChAT G21 to determine which structure is present.

4.3 Nitroxide Radical Synthesis

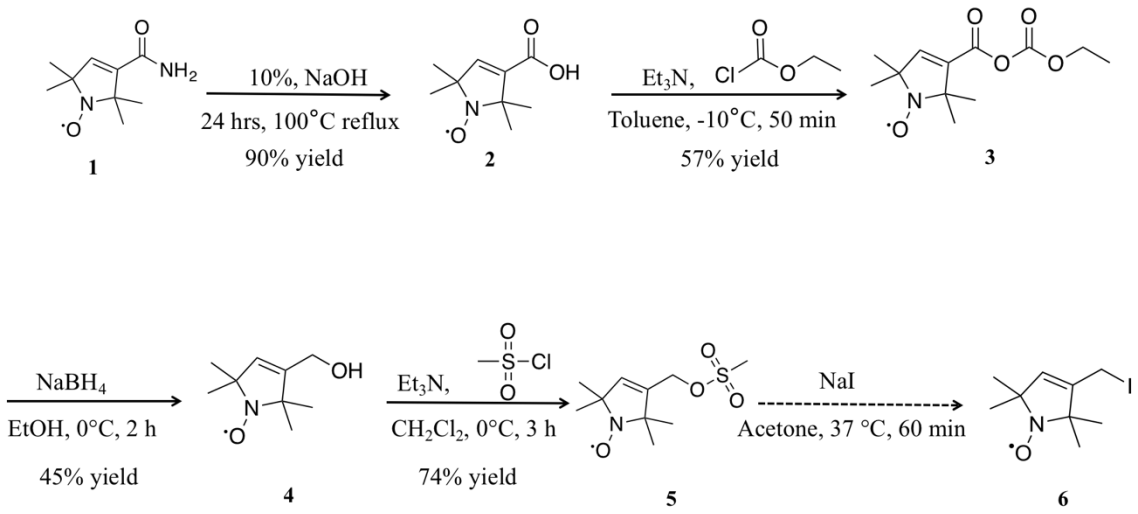

Figure 13. Synthesis of Nitroxide Spin Label.

For the analysis of ChAT G21 using EPR, the alkyl iodide nitroxide radical, compound 6, had to first be synthesized which can be seen in figure 13. Compound 1 was chosen as a starting point due to its commercial availability. The carbamide of compound 1 was hydrolyzed to produce the carboxylic acid, compound 2.[27] Compound 2 could have been reduced to compound 4 using $LiAlBH_4$ but this approach was not taken as it was reported in the literature that it resulted in low yield and a lack of reproducibility.[24] Ethylchloroformate was then added to compound 2 to create the anhydride compound 3.[24] The anhydride was then reduced to produce compound 4.[24] The methylsulfonyl moiety was added to compound 4 creating compound 5.[26]

The synthesis of compound 6 was unsuccessful but there was much learned that could help future studies. One of the problems observed with the synthesis was the

instability of the methylsulfonyl moiety of compound 5. The moiety made the compounds sensitive to light and also to heat resulting in its rapid decomposition. Compound 6 was reported as being even more unstable than compound 5 by Qin et al. Therefore, the synthesis of compound 6 was not performed due to insufficient amounts of purified compound 5. Compound 5 needed to be purified since the purification of compound 6 was not advised. This would ultimately reduce the amount of untagged DNA molecules for the EPR assays. Future experiments are needed to scale up compound 4 so that compound 5 could be produced as needed. Despite the harsh synthesis conditions, the protonation of the nitroxide radical was not observed. This was supported by the NMR radical shifting observed in synthesized compounds. This was further supported by comparing compound 4 to compound 1 using the EPR spectrophotometer. When analyzed in the EPR spectrophotometer no reduction in signal intensity was observed.

4.4 Future EPR Work

For EPR analysis of ChAT G21, more of compound 4 is needed to produce compound 6. The spin labels can then be added as described in Qin, et al. To distinguish which structure is present the tagged nucleotides will have to be chosen carefully. In figure 14, two nucleotides, on opposite strands of a B-DNA simulated ChAT G-21, are separated by 22.329 angstroms. The 25 nm max effective range of the CW-EPR distance measurement, is almost reached when the tagged nucleotides are within +/- 3 nucleotides of the nucleotide that is complementary to the spin label. In the parallel Q-GPX form the distances would be decreased due to its compact form and the distances separated the G runs would become less significant. In table 4, several ChAT G21 individual sequences

can be seen with phosphorothioate modifications. These sequences should allow for several combinations for analysis.

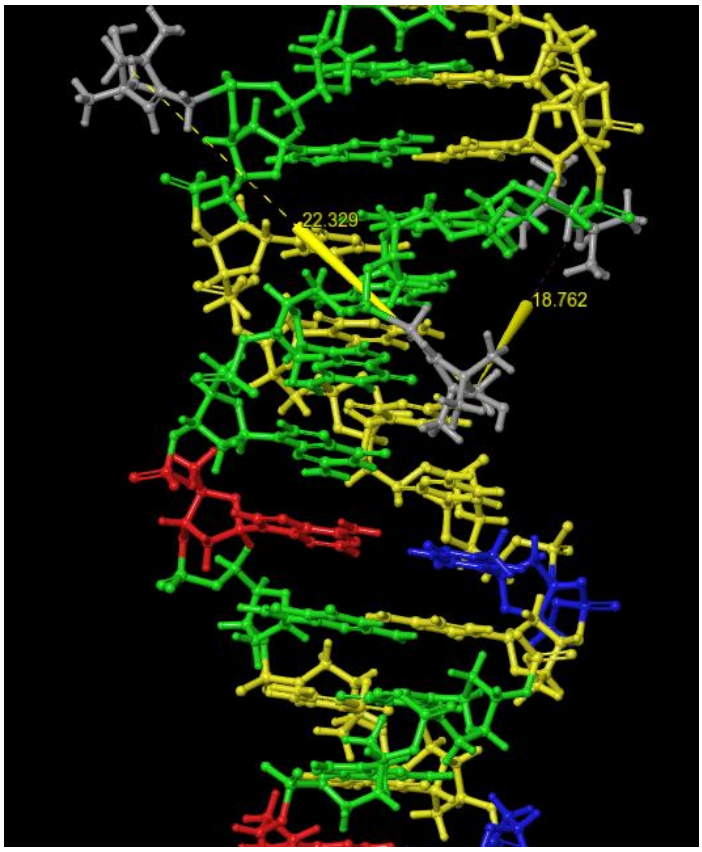

Figure 14. Simulated ChAT G21 B-DNA Structure with Nitroxide Radicals. A DNA structure was built using the biopolymer build tool of SybylX, a computational chemistry program. The positive strand contains the two nitroxide radicals while the negative strand contains the one nitroxide radical seen on the top left corner. The nitroxide radical seen in grey was built using Maestro a computational chemistry program. The attachment of the nitroxide radical to DNA ChAT G21 structure built on SybylX was done on Maestro, as well as the distance measurements. Nucleotides are depicted as such: guanine, green; cytosine, yellow; adenosine, red; and thymine, blue.

Table 4. Combination of ChAT G21 Single Strands with Phosphorothioate Modifications. This table demonstrates the phosphorothioate modification as *. The + and - signs refers to which strand of ChAT G21 the sequence belongs to.

Name	Sequence
ChAT G21+ mod 7	5'-CCCGGGG*GAGCCTGAGGACCC-3'
ChAT G21- mod 20	3'-GG*GCCCCCTCGGACTCCTGGG-5'
ChAT G21+ mod 7,17	5'-CCCGGGG*GAGCCTGAGG*ACCC-3'
ChAT G21-	3'-GGGCCCCCTCGGACTCCTGGG-5'
ChAT G21+ mod 2	3'-GGGCCCCCTCGGACTCCTG*GG-5'

Once the distances are taken, the distances can be compared using a molecular mechanics program. A possible template that can be used is the crystal structure of a telomeric bimolecular parallel G-QPX as a starting point for model building using SybylX or Maestro. The structure can then be built in a way that agrees with the distance measurements obtained from the EPR analysis. The models can also be further refined via energy minimization and molecular dynamics using the YASARA program.

4.5 Conclusion

Although the EPR distance measurement or mobility experiments could not be completed due to the nitroxide radical synthesis, the methodology for this experiment can be used in future studies. This work compiles the synthesis of compound 5 from commercially available and inexpensive reagents that can be found in a standard synthesis laboratory. Although there are several methods that use EPR to analyze G-

QPX, this work is the first in providing an EPR method that can potentially distinguish G-QPX structures from non-G-QPX structures.

BIBLIOGRAPHY

(1) Oda, Y. Choline Acetyltransferase: The Structure, Distribution and Pathologic Changes in the Central Nervous System. Pathol. Int. 1999, 49 (11), 921–937.

(2) Jones, C. K.; Byun, N.; Bubser, M. Muscarinic and Nicotinic Acetylcholine Receptor Agonists and Allosteric Modulators for the Treatment of Schizophrenia. Neuropsychopharmacology 2012, 37 (1), 16–42.

(3) Daly, J. W. Nicotinic Agonists, Antagonists, and Modulators from Natural Sources. Cell. Mol. Neurobiol. 2005, 25 (3–4), 513–552.

(4) Mendoza, O.; Bourdoncle, A.; Boulé, J. B.; Brosh, R. M.; Mergny, J. L. G-Quadruplexes and Helicases. Nucleic Acids Res. 2016, 44 (5), 1989–2006.

(5) Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M. Circular Dichroism and Conformational Polymorphism of DNA. Nucleic Acids Res. 2009, 37 (6), 1713–1725.

(6) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: Sequence, Topology and Structure. Nucleic Acids Res. 2006, 34 (19), 5402–5415.

(7) Adrian, M.; Heddi, B.; Phan, A. T. NMR Spectroscopy of G-Quadruplexes. Methods 2012, 57 (1), 11–24.

(8) Yang, L.; Qing, Z.; Liu, C.; Tang, Q.; Li, J.; Yang, S.; Zheng, J.; Yang, R.; Tan, W. Direct Fluorescent Detection of Blood Potassium by Ion-Selective Formation of Intermolecular G-Quadruplex and Ligand Binding. Anal. Chem. 2016, 88 (18), 9285–9292.

(9) Bedrat, A.; Lacroix, L.; Mergny, J. L. Re-Evaluation of G-Quadruplex Propensity with G4Hunter. Nucleic Acids Res. 2016, 44 (4), 1746–1759.

(10) Huppert, J. L.; Balasubramanian, S. Prevalence of Quadruplexes in the Human Genome. Nucleic Acids Res. 2005, 33 (9), 2908–2916.

(11) Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.; Balasubramanian, S. High-Throughput Sequencing of DNA G-Quadruplex Structures in the Human Genome. Nat. Biotechnol. 2015, 33 (8), 877–881.

(12) Balasubramanian, S.; Neidle, S. G-Quadruplex Nucleic Acids as Therapeutic Targets. Curr. Opin. Chem. Biol. 2009, 13 (3), 345–353.

(13) Bazzicalupi, C.; Ferraroni, M.; Bilia, A. R.; Scheggi, F.; Gratteri, P. The Crystal Structure of Human Telomeric DNA Complexed with Berberine: An Interesting Case of Stacked Ligand to G-Tetrad Ratio Higher than 1:1. Nucleic Acids Res. 2013, 41 (1), 632–638.

(14) Baghee Ravari, S. Ionic Modulation of QPX Stability as a Nano - Switch Regulating Gene Expression in Neurons, University of North Carolina at Greensboro, 2016.

(15) Mita, H.; Ohyama, T.; Tanaka, Y.; Yamamoto, Y. Formation of a Complex of 5,10,15,20-Tetrakis(N -Methylpyridinium-4-Yl)-21 H ,23 H -Porphyrin with G-Quadruplex DNA †. Biochemistry 2006, 45 (22), 6765–6772.

(16) Vorlíčková, M.; Kejnovská, I.; Bednářová, K.; Renčiuk, D.; Kypr, J. Circular Dichroism Spectroscopy of DNA: From Duplexes to Quadruplexes. Chirality 2012, 24 (9), 691–698.

(17) Qin, P. Z.; Haworth, I. S.; Cai, Q.; Kusnetzow, A. K.; Grant, G. P. G.; Price, E. A.; Sowa, G. Z.; Popova, A.; Herreros, B.; He, H. Measuring Nanometer Distances in Nucleic Acids Using a Sequence-Independent Nitroxide Probe. Nat. Protoc. 2007, 2 (10), 2354–2365.

(18) Kim, N.-K.; Murali, A.; DeRose, V. J. A Distance Ruler for RNA Using EPR and Site-Directed Spin Labeling. Chem. Biol. 2004, 11 (7), 939–948.

(19) Zhang, X.; Xu, C. X.; Di Felice, R.; Sponer, J.; Islam, B.; Stadlbauer, P.; Ding, Y.; Mao, L.; Mao, Z. W.; Qin, P. Z. Conformations of Human Telomeric G-Quadruplex Studied Using a Nucleotide-Independent Nitroxide Label. Biochemistry 2016, 55 (2), 360–372.

(20) Vorlíčková, M.; Kejnovská, I.; Sagi, J.; Renčiuk, D.; Bednářová, K.; Motlová, J.; Kypr, J. Circular Dichroism and Guanine Quadruplexes. Methods 2012, 57 (1), 64–75.

(21) Altona, C.; Sundaralingam, M. Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. a New Description Using the Concept of Pseudorotation. J. Am. Chem. Soc. 1972, 94 (23), 8205–8212.

(22) Stoll, S.; Schweiger, A. EasySpin, a Comprehensive Software Package for Spectral Simulation and Analysis in EPR. J. Magn. Reson. 2006, 178 (1), 42–55.

(23) Okamoto, A.; Inasaki, T.; Saito, I. Nitroxide-Labeled Guanine as an ESR Spin Probe for Structural Study of DNA. Bioorganic Med. Chem. Lett. 2004, 14 (13), 3415–3418.

(24) Hajjaj, B.; Shah, A.; Bell, S.; Shirran, S. L.; Botting, C. H.; Slawin, A. M. Z.; Hulme, A. N.; Lovett, J. E. Synthesis of Next-Generation Maleimide Radical Labels. Synlett 2016, 27 (16), 2357–2361.

(25) Grimaldi, M.; Scrima, M.; Esposito, C.; Vitiello, G.; Ramunno, A.; Limongelli, V.; D'Errico, G.; Novellino, E.; D'Ursi, A. M. Membrane Charge Dependent States of the β-Amyloid Fragment Aβ (16–35) with Differently Charged Micelle Aggregates. Biochim. Biophys. Acta - Biomembr. 2010, 1798 (3), 660–671.

(26) Haugland, M. M.; El-Sagheer, A. H.; Porter, R. J.; Peña, J.; Brown, T.; Anderson, E. A.; Lovett, J. E. 2′-Alkynylnucleotides: A Sequence- and Spin Label-Flexible Strategy for EPR Spectroscopy in DNA. J. Am. Chem. Soc. 2016, 138 (29), 9069–9072.

(27) Sato, S.; Tsunoda, M.; Suzuki, M.; Kutsuna, M.; Takido-uchi, K.; Shindo, M.; Mizuguchi, H.; Obara, H.; Ohya, H. Synthesis and Spectral Properties of Polymethine-Cyanine Dye-Nitroxide Radical Hybrid Compounds for Use as Fluorescence Probes to Monitor Reducing Species and Radicals. Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 2009, 71 (5), 2030–2039.

(28) Martinie, J.; Michon, J.; Rassat, A. Nitroxides. LXX. Electron Spin Resonance Study of Cyclodextrin Inclusion Compounds. J. Am. Chem. Soc. 1975, 97 (7), 1818–1823.

(29) Banham, J. E.; Baker, C. M.; Ceola, S.; Day, I. J.; Grant, G. H.; Groenen, E. J. J.; Rodgers, C. T.; Jeschke, G.; Timmel, C. R. Distance Measurements in the Borderline Region of Applicability of CW EPR and DEER: A Model Study on a Homologous Series of Spin-Labelled Peptides. J. Magn. Reson. 2008, 191 (2), 202–218.

(30) Mchaourab, H. S.; Oh, K. J.; Fang, C. J.; Hubbell, W. L. Conformation of T4 Lysozyme in Solution. Hinge-Bending Motion and the Substrate-Induced Conformational Transition Studied by Site-Directed Spin Labeling. Biochemistry 1997, 36 (2), 307–316.

(31) Cooke, J. A.; Brown, L. J. Distance Measurements by Continuous Wave EPR Spectroscopy to Monitor Protein Folding. In Protein Folding, Misfolding, and Disease: Methods and Protocols; Springer, 2011; Vol. 752, pp 73–96.

(32) Agrawal, P.; Lin, C.; Mathad, R. I.; Carver, M.; Yang, D. The Major G-Quadruplex Formed in the Human BCL-2 Proximal Promoter Adopts a Parallel Structure with a 13-Nt Loop in K+ Solution. J. Am. Chem. Soc. 2014, 136 (5), 1750–1753.